2020
ACTIVITY
REPORT

Project-Team
KERDATA

RESEARCH CENTRE

Rennes - Bretagne Atlantique

IN PARTNERSHIP WITH:

Institut national des sciences appliquées de Rennes, Université Rennes 1, École normale supérieure de Rennes

Scalable Storage for Clouds and Beyond

IN COLLABORATION WITH: Institut de recherche en informatique et systèmes aléatoires (IRISA)

DOMAIN

Networks, Systems and Services, Distributed Computing

THEME

Distributed and High Performance Computing

Contents

Project-Team KERDATA

Creation of the Team: 2009 July 01, updated into Project-Team: 2012 July 01

Keywords

Computer sciences and digital sciences

A1.1.1. – Multicore, Manycore

A1.1.4. – High performance computing

A1.1.5. – Exascale

A1.1.9. – Fault tolerant systems

A1.3. – Distributed Systems

A1.3.5. – Cloud

A1.3.6. – Fog, Edge

A1.6. – Green Computing

A2.6.2. – Middleware

A3.1.2. – Data management, quering and storage

A3.1.3. – Distributed data

A3.1.8. – Big data (production, storage, transfer)

A6.2.7. – High performance computing

A6.3. – Computation-data interaction

A7.1.1. – Distributed algorithms

A9.2. – Machine learning

A9.7. – AI algorithmics

Other research topics and application domains

B3.2. – Climate and meteorology

B3.3.1. – Earth and subsoil

B8.2. – Connected city

B9.5.6. – Data science

B9.8. – Reproducibility

B9.11.1. – Environmental risks

1 Team members, visitors, external collaborators

Research Scientists

- Gabriel Antoniu [Team leader, Inria, Senior Researcher, HDR]

- François Tessier [Inria, from Nov 2020, Starting Faculty Position]

Faculty Members

- Luc Bougé [École normale supérieure de Rennes, Professor, HDR]

- Alexandru Costan [INSA Rennes, Associate Professor, HDR]

PhD Student

- Daniel Rosendo [Inria]

Technical Staff

- Thomas Bouvier [Inria, From January 2020, Attached to STIP, activity fully dedicated to KerData (software development for KerA)]

- Joshua Charles Bowden [Inria, Engineer, from Apr 2020]

Administrative Assistant

- Gaelle Tworkowski [Inria]

2 Overall objectives

2.1 Context: the need for scalable data management

We are witnessing a rapidly increasing number of application areas generating and processing very large volumes of data on a regular basis. Such applications are called *data-intensive*. Governmental and commercial statistics, climate modeling, cosmology, genetics, bio-informatics, high-energy physics are just a few examples in the scientific area. In addition, rapidly growing amounts of data from social networks and commercial applications are now routinely processed.

In all these examples, the overall application performance is highly dependent on the properties of the underlying data management service. It becomes crucial to store and manipulate massive data efficiently. However, these data are typically *shared* at a large scale and *concurrently accessed* at a high degree. With the emergence of infrastructures such as cloud and edge computing platforms and post-Petascale high-performance computing (HPC) systems, achieving highly scalable data management under such conditions has become a major challenge.

2.2 Our objective

The KerData project-team is namely focusing on designing innovative architectures and systems for *scalable data storage and processing*. We target three types of infrastructures: *edge*, *clouds* and *post-Petascale high-performance supercomputers*, according to the current needs and requirements of data-intensive applications.

We are especially concerned by the applications of major international and industrial players in cloud computing and extreme-scale high-performance computing (HPC), which shape the long-term agenda of the cloud computing [29, 25] and Exascale HPC [28] research communities. The Big Data area emphasized the challenges related to Volume, Velocity and Variety. This is yet another element of context that further highlights the primary importance of designing data management systems that are efficient at a very large scale.

Alignment with Inria's scientific strategy   Data-intensive applications exhibit several common requirements with respect to the need for data storage and I/O processing. We focus on some core challenges related to data management, resulted from these requirements. Our choice is perfectly in line with Inria's strategic objectives [32], which acknowledges HPC-Big Data convergence as one of the Top 3 priorities of our institute.

In the area of cloud data processing, a significant milestone is the emergence of the Map-Reduce [20] parallel programming paradigm. It is currently used on most cloud platforms, following the trend set up by Amazon [22]. At the core of Map-Reduce frameworks lies the storage system, a key component which must meet a series of specific requirements that are not fully met yet by existing solutions: the ability to provide efficient *fine-grain access* to the files, while sustaining a *high throughput* in spite of *heavy access concurrency*; the need to provide a high resilience to *failures*; the need to take *energy-efficiency* issues into account.

More recently, it becomes clear that data-intensive processing needs to go beyond the frontiers of a single type of infrastructure. Cloud workflows evolve from single-datacenter deployment towards multiple-datacenter deployments, and further from cloud deployments towards distributed, edge-based infrastructures. In this perspective, extra challenges arise, related to the efficiency of metadata management and data processing.

Challenges and goals related to data-intensive HPC applications   Key research fields such as climate modeling, solid Earth sciences or astrophysics rely on very large-scale simulations running on post-Petascale supercomputers. Such applications exhibit requirements clearly identified by international panels of experts like IESP [31], EESI [27], ETP4HPC [28]. A jump of one order of magnitude in the size of numerical simulations is required to address some of the fundamental questions in several communities in this context. In particular, the lack of data-intensive infrastructures and methodologies to analyze the huge results of such simulations is a major limiting factor.

The challenge we have been addressing is to find new ways to store, visualize and analyze massive outputs of data during and after the simulations. Our main initial goal was to do it without impacting the overall performance, avoiding the *jitter* generated by I/O interference as much as possible. During the recent years we focused specifically on *in situ processing* approaches and we explored approaches to *model and predict I/O phase occurrences* and to *reduce intra-application and cross-application I/O interference*. More recently, we started to investigate approaches exploring how to process data for complex workflows where simulations run on HPC systems are coupled with data analytics codes typically running on clouds. A particularly challenging case is that of such applications running on hybrid HPC/cloud/edge infrastructures.

2.3   Our approach

KerData's global approach consists in studying, designing, implementing and evaluating distributed algorithms and software architectures for scalable data storage and I/O management for efficient, large-scale data processing. We target three main execution infrastructures: edge and cloud platforms and post-Petascale HPC supercomputers.

Platforms and Methodology   The highly experimental nature of our research validation methodology should be emphasized. To validate our proposed algorithms and architectures, we build

software prototypes, then validate them at a large scale on real testbeds and experimental platforms.

We strongly rely on the Grid'5000 platform. Moreover, thanks to our projects and partnerships, we have access to reference software and physical infrastructures. In the cloud area, we use the Microsoft Azure, Amazon cloud platforms and the Chameleon [23] experimental cloud testbed. In the post-Petascale HPC area, we are running our experiments on systems including some top-ranked supercomputers, such as Titan, Jaguar, Kraken or Aurora. This provides us with excellent opportunities to validate our results on advanced realistic platforms.

Collaboration strategy   Our collaboration portfolio includes international teams that are active in the areas of data management for edge, clouds and HPC systems, both in Academia and Industry.

Our academic collaborating partners include Argonne National Lab, University of Illinois at Urbana-Champaign, Universidad Politécnica de Madrid, Barcelona Supercomputing Center. In industry, through bilateral or multilateral projects, we have been collaborating with Microsoft, IBM, Total, Huawei, ATOS/BULL.

Moreover, the consortiums of our collaborative projects include application partners in multiple application domains from the areas of area of climate modeling, precision agriculture, earth sciences, precision agriculture, smart cities, botanical science. This multidisciplinary approach is an additional asset, which enables us to take into account application requirements in the early design phase of our proposed approaches to data storage and processing, and to validate those solutions with real applications and real users.

3   Research program

3.1   Research axis 1: Convergence of HPC and Big Data

The tools and cultures of High Performance Computing and Big Data Analytics have evolved in divergent ways. This is to the detriment of both. However, big computations still generate and are needed to analyze Big Data. As scientific research increasingly depends on both high-speed computing and data analytics, the potential interoperability and scaling convergence of these two ecosystems is crucial to the future.

Our objective is premised on the idea that we must explore the ways in which the major challenges associated with Big Data analytics intersect with, impact, and potentially change the directions now in progress for achieving Exascale computing.

In particular, a key milestone is to achieve convergence through common abstractions and techniques for data storage and processing in support of complex workflows combining simulations and analytics. Such application workflows need such a convergence to run on hybrid infrastructures combining HPC systems and clouds (potentially extended to edge devices, in a complete digital continuum).

> | Collaboration. | *This axis is addressed in close collaboration with María Pérez (UPM), Rob Ross (ANL), Toni Cortes (BSC), Several groups at Argonne National Laboratory and NCSA (Franck Cappello, Rob Ross, Bill Kramer, Tom Peterka).*
>
> *Relevant groups with similar interests are the following ones.*
>
> - *The group of Jack Dongarra, Innovative Computing Laboratory at University of Tennessee, who is leading international efforts for the convergence of Exascale Computing and Big Data.*
> - *The group of Satoshi Matsuoka, RIKEN, working on system software for clouds and HPC.*
> - *The group of Ian Foster, Argonne National Laboratory, working on on-demand data analytics and storage for extreme-scale simulations and experiments.*

3.1.1   Towards common storage abstractions for Extreme Computing and Big Data analytics

Storage is a plausible pathway to convergence. In this context, we plan to focus on the needs of concurrent Big Data applications that require high-performance storage, as well as transaction support. Although blobs (binary large objects) are an increasingly popular storage model for such applications, state-of-the-art blob storage systems offer no transaction semantics. This demands users to coordinate data access carefully in order to avoid race conditions, inconsistent writes, overwrites and other problems that cause erratic behavior.

There is a gap between existing storage solutions and application requirements, which limits the design of transaction-oriented applications. In this context, one idea on which we focus our efforts is exploring how blob storage systems could provide built-in, multiblob transactions, while retaining sequential consistency and high throughput under heavy access concurrency. From a more general perspective, we investigate how object-based storage could serve as a common storage abstraction for both HPC simulations (traditionally running on supercomputers) and for Big Data analytics codes (traditionally running on clouds). HPC/cloud storage convergence can be seen as as a means to make a step forward towards the general HPC-Big Data convergence.

The early principles of this research direction have already raised interest from our partners at ANL (Rob Ross) and UPM (María Pérez) for potential collaborations. Our paper on the Týr transactional blob storage system, selected as a Best Student Paper Award Finalist at the SC16 conference [9], illustrates this direction.

3.1.2   Towards unified data processing techniques for Extreme Computing and Big Data applications

In the high-performance computing area (HPC), the need to get fast and relevant insights from massive amounts of data generated by extreme-scale computations led to the emergence of *in situ processing*. It allows data to be visualized and processed in real-time on the supercomputer generating them, in an interactive way, as they are produced, as opposed to the traditional approach consisting of transferring data off-site after the end of the computation, for offline analysis. As such processing runs on the same resources executing the simulation, there is a risk to "disturb" the simulation if it consumes too many resources.

Consequently, an alternative approach was proposed (*in transit processing*), as a means to reduce this impact: data are transferred to some temporary processing resources (with high memory and processing capacities). After this real-time processing, they are moved to persistent storage.

In the Big Data area, the search for real-time, fast analysis was materialized through a different approach: stream-based processing. Such an approach is based on a different abstraction for data, that are seen as a dynamic flow of items to be processed. Stream-based processing and in situ/in transit processing have been developed separately and implemented in different tools in the BDA (*Big Data Applications*) and HPC (*High-Performance Computing*) areas respectively.

A major challenge from the perspective of the HPC-BDA convergence is their joint use in a unified data processing architecture. This is one of the future research challenges that we plan to address in the near future, by combining ongoing approaches currently active in our team: Damaris and KerA. We started preliminary work within the *"Frameworks"* work package of the HPC-Big Data IPL. Further exploring this convergence is a core direction of our involvement in collaborative European projects: PRACE 6IP (2019-2021); and ACROSS, to start in 2021.

3.2   Research axis 2: Cloud and Edge processing

The recent evolutions in the area of Big Data processing have pointed out some limitations of the initial Map-Reduce model. It is well suited for batch data processing, but less suited for real-time processing of dynamic data streams. New types of data-intensive applications emerge, e.g., for enterprises who need to perform analysis on their stream data in ways that can give fast

results (i.e., in real time) at scale (e.g., click-stream analysis and network-monitoring log analysis). Similarly, scientists require fast and accurate data processing techniques in order to analyze their experimental data correctly at scale (e.g., collective analysis of large data sets distributed in multiple geographically distributed locations).

Our plan is to revisit current data storage and processing techniques to cope with the volatile requirements of data-intensive applications on large-scale dynamic clouds in a cost-efficient way, with a particular focus on streaming. More recently, the strong emergence of edge/fog-based infrastructures leads to additional challenges for new scenarios involving hybrid cloud/fog/edge systems.

> **Collaboration.** *This axis is addressed in close collaboration with María Pérez (UPM), Kate Keahey (ANL).*
>
> *Relevant groups with similar interests include the following ones.*
>
> - *The group of Geoffrey Fox, Indiana University, working on data analytics, cloud data processing, stream processing.*
> - *The group at RISE Lab, UC Berkeley, working on real-time stream-based processing and analytics.*
> - *The group of Ewa Deelman, USC Information Sciences Institute, working on resource management for workflows in clouds.*

3.2.1   Stream-oriented, Big Data processing on clouds

The state-of-the-art Hadoop Map-Reduce framework cannot deal with stream data applications, as it requires the data to be initially stored in a distributed file system in order to process them. To better cope with the above-mentioned requirements, several systems have been introduced for stream data processing such as Flink [30], Spark [33], Storm [34], and Google MillWheel [18]. These systems keep active data in memory to decrease latency, and preserve scalability by using partitioning data over the various distributed nodes or dividing the streams into a set of deterministic batch computations to be distributedly processed.

However, they are designed to work in dedicated environments and they do not consider the performance variability (i.e., network, I/O, etc.) caused by resource contention in the cloud. This variability may in turn cause high and unpredictable latency when output streams are transmitted to further analysis. Moreover, they overlook the dynamic nature of data streams and the volatility in their computation requirements. Finally, they still address failures in a best-effort manner. Our objective is to investigate new approaches for reliable, stream Big Data processing on clouds.

3.2.2   Efficient edge, cloud and hybrid edge/cloud data processing

Today, we are approaching an important technological milestone: applications are generating huge amounts of data and are demanding low-latency responses to their requests. Mobile computing and Internet of Things (IoT) applications are good illustrations of such scenarios. Using only cloud computing for such scenarios is challenging. Firstly, cloud resources are most of the time accessed through Internet, hence, data are sent across high-latency wide area networks, which may degrade the performance of applications. Secondly, it may be impossible to send data to the cloud due to data regulations, national security laws or simply because an Internet connection is not available. Finally, data transmission costs (e.g., cloud provider fees, carrier costs) could make a business solution impractical.

Edge computing is a new paradigm which aims to address some of these issues. The key idea is to leverage computing and storage resources at the "edge" of the network, i.e., on processing units located close to the data sources. This allows applications to outsource task execution from the main (cloud) processing data centers to the edge. The development of edge computing was accelerated by the recent emergence of stream processing, a new model for handling continuous

flows of data in real-time, as opposed to batch processing, which typically processes bounded datasets offline.

However, edge computing is not a silver bullet: issues like node volatility, limited processing power, high latency between nodes, fault tolerance and data degradation may impact applications depending on the characteristics of the infrastructure.

Some relevant research questions are: How much can one improve (or degrade) the performance of an application by performing data processing closer to the data sources rather than performing it in the cloud? How to progress towards a seamless scheduling and execution of a data analytics workflow and break the limitations of the current dual approaches? Those dual approached used to be applied in preliminary efforts in this area, which rely on manual and empirical deployment of the corresponding dataflow operator graphs, using separate analytics engines for centralized clouds and for edge systems respectively.

Our objective is to try to precisely answer such questions. We are interested in understanding the conditions that enable the usage of edge or cloud computing to reduce the time to results and the associated costs. While some state-of-the-art approaches advocate either "100 % cloud" or "100 % edge" solutions, the relative efficiency of a method over the other may vary. Intuitively, it depends on many parameters, including network technology, hardware characteristics, volume of data, amount of computing power, processing framework configuration and application requirements, to cite a few. We plan to study their impact on the overall application performance.

3.3 Research axis 3: Supporting AI across the digital continuum

Integrating and processing high-frequency data streams from multiple sensors scattered over a large territory in a timely manner requires high-performance computing techniques and equipments. For instance, a machine learning earthquake detection solution has to be designed jointly with experts in distributed computing and cyber-infrastructure to enable real-time alerts. Because of the large number of sensors and their high sampling rate, a traditional centralized approach which transfers all data to a single point may be impractical. Our goal is to investigate innovative solutions for the design of efficient data processing infrastructures for a distributed machine learning-based approach.

In particular, building on our previous results in the area of efficient stream processing systems, we aim to explore approaches for unified data storage, processing and machine-learning based analytics across the whole digital continuum (i.e., for highly distributed applications deployed on hybrid edge/cloud/HPC infrastructures).

Collaboration. *This recently started axis is worked out in close collaboration with the group of Manish Parashar, Rutgers University, and with the LACODAM team at Inria, focused on large-scale collaborative data mining.*

This work led to a joint paper published at AAAI-20, a reference A* conference in the area of Artificial Intelligence, where it was distinguished with an Outstanding Paper Award - Special Track for Social Impact [15].

4 Application domains

The KerData team investigates the design and implementation of architectures for data storage and processing across clouds, HPC and edge-based systems, which address the needs of a large spectrum of applications. The use cases we target to validate our research results come from the following domains.

4.1 Climate and meteorology

The European Centre for Medium-Range Weather Forecasts (ECMWF) [26] is one of the largest weather forecasting centers in the world that provides data to national institutions and private clients. ECMWF's production workflow collects data at the edge through a large set of sensors (satellite devices, ground sensors, smart sensors). This data, approximately 80 millions of observations per day, is then moved to be assimilated, i.e. analyzed and sorted, before being sent to a supercomputer to feed the prediction models.

The compute and I/O intensive large-scale simulations built upon these models use ensemble forecasting methods for the refinement. To date, these simulations generate approximately 60 TB per hour, while the center predicts an annual increase of 40 % of this volume. Structured datasets called "products" are then generated from this output data and are disseminated to different clients, such as public institutions or private companies, at a rate of 1PB per month transmitted.

In the framework of the ACROSS EuroHPC Project which has been accepted in 2020 (See 9.3.1), our goal is to participate in the design of a hybrid software stack for the HPC, Big Data and AI domains. This software stack must be compatible with a wide range of heterogeneous hardware technologies and must meet the needs of the trans-continuum ECMWF workflow.

4.2 Earth science

Earthquakes cause substantial loss of life and damage to the built environment across areas spanning hundreds of kilometers from their origins. These large ground motions often lead to hazards such as tsunamis, fires and landslides. To mitigate the disastrous effects, a number of Earthquake Early Warning (EEW) systems have been built around the world. Those critical systems, operating 24/7, are expected to automatically detect and characterize earthquakes as they happen, and to deliver alerts before the ground motion actually reaches sensitive areas so that protective measures could be taken.

Our research aims to improve the accuracy of Earthquake Early Warning (EEW) systems. These systems are designed to detect and characterize medium and large earthquakes before their damaging effects reach a certain location. Traditional EEW methods based on seismometers fail to accurately identify large earthquakes due to their low sensitivity to ground motion velocity. The recently introduced high-precision GPS stations, on the other hand, are ineffective to identify medium earthquakes due to their propensity to produce noisy data. In addition, GPS stations and seismometers may be deployed in large numbers across different locations and may produce a significant volume of data consequently, affecting the response time and the robustness of EEW systems.

Integrating and processing in a timely manner high-frequency data streams from multiple sensors scattered over a large territory requires high-performance computing techniques and equipments. We therefore design distributed machine learning-based approaches [15] to earthquake detection, jointly with experts in machine learning and Earth data. Our expertise in swift processing of data on edge and cloud infrastructures allows to learn from the data from the large number of sensors arriving at high sampling rate, without transferring all data to a single point and thus enables real-time alerts.

4.3 Sustainable development through precision agriculture

Feeding the growing world's population is a decisive challenge, especially in view of climate change, which adds a certain level of uncertainty in food production. Sustainable and precision agriculture is one of the answers that can be implemented to partly overcome this issue. Precision agriculture consists in using new technologies to improve crop management by considering environmental parameters such as temperature, soil moisture or weather conditions, for example. These techniques now need to scale up to improve their accuracy. Since a few years, we have seen

the emergence of precision agriculture workflows running across the digital continuum, that is to say all the computing resources from the edge to High-Performance Computing (HPC) and Cloud-type infrastructures. This move to scale is accompanied by new problems, particularly with regard to data movements.

CybeleTech [24] is a French company that aims at developing the use of numerical technologies in agriculture. The core products of CybeleTech are based on numerical simulation of plant growth through dedicated biophysical models and machine learning methods extracting knowledge through large databases. To feed its models, CybeleTech collects data from sensors installed on open agricultural plots or in crop greenhouses. Plant growth models take weather variables as input and the accuracy of agronomic indices estimation heavily rely on the accuracy of these variables.

To this purpose, CybeleTech wishes to collect precise meteorological information from large forecasting centers such as the European Center for Medium-Range Weather Forecasting (ECMWF) [26]. This data gathering is not trivial since it involves large data movements between two distant sites under severe time constraints. Our objective here in the context of the EUPEX EuroHPC project (submitted, under review. See 9.3.1) is to propose new data management techniques and data movement algorithms to accelerate the execution of these hybrid geo-distributed workflows running on large-scale systems in the area of precision agriculture.

4.4   Smart cities

The proliferation of small sensors and devices that are capable of generating valuable information in the context of the Internet of Things (IoT) has exacerbated the amount of data flowing from all connected objects to cloud infrastructures. In particular, this is true for Smart City applications. These applications raise specific challenges, as they typically have to handle small data (in the order of bytes and kilobytes), arriving at high rates, from many geographical distributed sources (sensors, citizens, public open data sources, etc.) and in heterogeneous formats, that need to be processed and acted upon with high reactivity in near real-time.

Our vision is that, by smartly and efficiently combining the data-driven analytics at the edge and in the cloud, it becomes possible to make a substantial step beyond state-of-the-art prescriptive analytics through a new, high-potential, faster approach to react to the sensed data of the smart cities. The goal is to build a data management platform that will enable comprehensive joint analytics of past (historical) and present (real-time) data, in the cloud and at the edge, respectively, allowing to quickly detect and react to special conditions and to predict how the targeted system would behave in critical situations. This vision is the driving objective of our SmartFastData associate team with Instituto Politécnico Nacional, Mexico.

4.5   Botanical Science

Pl@ntNet [21] is a large-scale participatory platform dedicated to the production of botanical data through AI-based plant identification. Pl@ntNet's main feature is a mobile app allowing smartphone owners to identify plants from photos and share their observations. It is used by around 10 million users all around the world (more than 180 countries) and it processes about 400K plant images per day. One of the challenges faced by Pl@ntNet engineers is to anticipate what should be the appropriate evolution of the infrastructure to pass the next spring peak without problems and also to know what should be done the following years.

Our research aims to improve the performance of Pl@ntNet. Reproducible evaluations of Pl@ntNet on large-scale testbed (e.g., deployed on Grid'5000 [19] by E2Clab [16]), aim to optimize its software configurations in order to minimize the user response time.

5 Social and environmental responsibility

5.1 Footprint of research activities

HPC facilities are expensive in capital outlay (both monetary and human) and in energy use. Our work on Damaris supports the efficient use of high performance computing resources. Damaris [2] can help minimize power needed in running computationally demanding engineering applications and can reduce the amount of storage used for results, thus supporting environmental goals and improving the cost effectiveness of running HPC systems.

5.2 Impact of research results

Social impact. One of our target applications is Early Earthquake Warning. We propose a solution that enables earthquakes classification with an outstandingly perfect accuracy. By enabling accurate identification of strong earthquakes, it becomes possible to trigger adequate measures and save lifes. For this reason, our work was distinguished with an Outstanding Paper Award — Special Track for Social Impact at AAAI-20, an A* conference in the area of Artificial Intelligence. This result has been highlighted by the *Le Monde* journal in its edition of December 28, 2020, in a section entitled: *Ces découvertes scientifiques que le Covid-19 a masquées en 2020*.

Environmental impact. As presented in Section 4, we have started a collaboration with Cy-beleTech that we plan to materialize in the framework of the EUPEX EuroHPC project. CybeleTech is a French company specialized in precision agriculture. Within the framework of our collaboration, we propose to focus our efforts on a scale-oriented data management mechanism targeting two CybeleTech use-cases. They address irrigation scheduling for orchards and optimal harvest date for corn, and their models require the acquisition of large volumes of remote data. The results of our collaboration will have concrete applications as they will improve the accuracy of plant growth models and improve decision making for precision agriculture, which directly aims to contribute to sustainable development.

6 Highlights of the year

6.1 Awards

Our work on a distributed approach to machine learning applied to early earthquake warnings [15] was distinguished through an Outstanding Paper Award - Special Track for Social Impact at AAAI-20, an A* conference in the area of Artificial Intelligence. This work has been highlighted by the *Le Monde* journal in its edition of December 28, 2020: *Ces découvertes scientifiques que le Covid-19 a masquées en 2020*.

6.2 New H2020 Project accepted

In the area of HPC-Big Data-AI convergence, a major achievement in 2020 is the acceptance of **a new European EuroHPC project**: H2020-JTI-EuroHPC ACROSS (2021-2023). KerData is coordinating Inria's contribution to this project, which includes a transfer action for the Damaris technology.

7   New software and platforms

7.1   New software

7.1.1   Damaris

Keywords: Visualization, I/O, HPC, Exascale, High performance computing

Scientific Description: Damaris is a middleware for I/O and data management targeting large-scale, MPI-based HPC simulations. It initially proposed to dedicate cores for asynchronous I/O in multicore nodes of recent HPC platforms, with an emphasis on ease of integration in existing simulations, efficient resource usage (with the use of shared memory) and simplicity of extension through plug-ins. Over the years, Damaris has evolved into a more elaborate system, providing the possibility to use dedicated cores or dedicated nodes to in situ data processing and visualization. It proposes a seamless connection to the VisIt visualization framework to enable in situ visualization with minimum impact on run time. Damaris provides an extremely simple API and can be easily integrated into the existing large-scale simulations.

Damaris was at the core of the PhD thesis of Matthieu Dorier, who received an Accessit to the Gilles Kahn Ph.D. Thesis Award of the SIF and the Academy of Science in 2015. Developed in the framework of our collaboration with the JLESC – Joint Laboratory for Extreme-Scale Computing, Damaris was the first software resulted from this joint lab validated in 2011 for integration to the Blue Waters supercomputer project. It scaled up to 16,000 cores on Oak Ridge's leadership supercomputer Titan (first in the Top500 supercomputer list in 2013) before being validated on other top supercomputers. Active development is currently continuing within the KerData team at Inria, where it is at the center of several collaborations with industry as well as with national and international academic partners.

Functional Description: Damaris is a middleware for data management and in-situ visualization targeting large-scale HPC simulations: - In situ data analysis by some dedicated cores/nodes of the simulation platform - Asynchronous and fast data transfer from HPC simulations to Damaris - Semantic-aware dataset processing through Damaris plug-ins - Writing aggregated data (by hdf5 format) or visualizating them either by VisIt or ParaView

URL: https://project.inria.fr/damaris/

Authors: Matthieu Dorier, Gabriel Antoniu, Orçun Yildiz, Luc Bougé, Hadi Salimi, Joshua Charles Bowden

Contact: Gabriel Antoniu

Participants: Gabriel Antoniu, Lokman Rahmani, Luc Bougé, Matthieu Dorier, Orçun Yildiz, Hadi Salimi, Joshua Charles Bowden

Partner: ENS Rennes

7.1.2   KerA

Name: KerAnalytics

Keywords: Big data, Distributed Storage Systems, Streaming, Real time

Scientific Description: Current state-of-the-art Big Data analytics architectures are built on top of a three layer stack: data streams are first acquired by the ingestion layer (e.g., Kafka) and then they flow through the processing layer (e.g., Flink) which relies on the storage layer (e.g., HDFS) for storing aggregated data or for archiving streams for later processing. Unfortunately, in spite of potential benefits brought by specialized layers (e.g., simplified

implementation), moving large quantities of data through specialized layers is not efficient: instead, data should be acquired, processed and stored while minimizing the number of copies.

KerA was developed on top of RAMCloud, a low-latency key-value distributed system.

Functional Description: KerA is a platform for continuous data ingestion and storage. This data can be of various types (videos, images, texts, structured data, etc.), and can be generated by multiple sources: sensors, websites, complex software (producers). The data thus acquired are stored permanently and made accessible to processing software (consumers), useful for extracting useful information or calculating statistics.

KerA focuses on reducing the latency between reading and publishing the data: the results are close to being obtained in real time. The platform is distributed: the available hardware resources are not on the same machine. This feature ensures scalability when the number of producers/consumers is increased and facilitates the replication of data on several servers in order to tolerate possible failures of some of them.

URL: https://kerdata.gitlabpages.inria.fr/Kerdata-Codes/kera-website/

Publications: hal-01773799, tel-02127065, hal-01532070

Authors: Ovidiu-Cristian Marcu, Gabriel Antoniu, Alexandru Costan, Thomas Bouvier

Contacts: Thomas Bouvier, Gabriel Antoniu, Alexandru Costan

7.1.3   Pufferscale

Keywords: Distributed Storage Systems, Elasticity

Scientific Description: Pufferscale is a rescaling service developed in the context of the Mochi project. It is a library designed to manage data units (buckets) of behalf of microservices during rescaling operations (adding or removing nodes). It tracks the position of each bucket in the system, and during rescaling operations Pufferscale determines the new position of each bucket (moved out of removed nodes, added to newly added nodes to balance the load). It gives the transfer instructions to the microservices.

Functional Description: Pufferscale is a rescaling service designed to manage data units (buckets) of behalf of microservices during rescaling operations (adding or removing nodes). It was design in the context of the Mochi project.

Release Contributions: First release of Pufferscale.

Contacts: Nathanaël Cheriere, Gabriel Antoniu

7.1.4   E2Clab

Name: Edge-to-Cloud lab

Keywords: Distributed Applications, Distributed systems, Computing Continuum, Large scale, Experimentation, Evaluation, Reproducibility

Functional Description: E2Clab allows researchers to reproduce in a representative way the application behavior in a controlled environment for extensive experiments and therefore to understand end-to-end performance of applications by correlating results to the parameter settings. E2Clab provides a rigorous approach to answering questions like: How to identify infrastructure bottlenecks? Which system parameters and infrastructure configurations impact on performance and how?

High-level features provided by E2Clab: - Reproducible Experiments: Supports repeatability, replicability and reproducibility. - Mapping: Application parts (Edge, Fog and Cloud/HPC) and physical testbed. - Variation & Scaling: Experiment variation and transparent scaling of scenarios. - Network Emulation: Edge-to-Cloud communication constraints. - Experiment Management: Deployment, execution and monitoring (e.g. on Grid'5000).

URL: https://kerdata.gitlabpages.inria.fr/Kerdata-Codes/e2clab/

Contacts: Daniel Rosendo, Gabriel Antoniu, Alexandru Costan, Matthieu Simonin

8 New results

8.1 Convergence HPC and Big Data

8.1.1 Unifying HPC and Big Data stacks: Towards application-defined blobs at the storage layer

Participants Alexandru Costan, Gabriel Antoniu.

| **Collaboration.** | *This work has been carried out in close co-operation with Pierre Matri, formerly a PhD intern in the team, and now at Argonne National Laboratory, USA.* |

HPC and Big Data stacks evolved separately. The storage layer offers opportunities for convergence, as the challenges associated with HPC and Big Data storage are similar: trading versatility for performance. This motivates a global move towards dropping file-based, POSIX-IO compliance systems.

However, on HPC platforms this is made difficult by the centralized storage architecture using file-based storage. In [13] we advocate that the growing trend of equipping HPC compute nodes with local storage redistributes the cards by enabling object storage to be deployed alongside the application on the compute nodes. Such integration of application and storage not only allows fine-grained configuration of the storage system, but also improves application portability across platforms. In addition, the single-user nature of such application-specific storage obviates the need for resource-consuming storage features like permissions or file hierarchies offered by traditional file systems.

We propose and evaluate Blobs (Binary Large Objects) as an alternative to distributed file systems. We factually demonstrate that it offers drop-in compatibility with a variety of existing applications while improving storage throughput by up to 28%.

8.1.2 Memory and data-awareness in hybrid workflows

Participant François Tessier.

Among the broad variety of challenges that arise from HPC and HPDA workloads, data movement is of paramount importance, especially on coming Exascale systems featuring multiple tiers of memory and storage. While the focus has, for years, been primarily on optimizing floating point operations, the importance of improving data handling on such architectures is now well understood. As optimization techniques can be applied at different stages (operating system, runtime system, programming environment, and so on), a middleware providing a uniform and consistent data-awareness becomes necessary.

We introduce a novel memory- and data-aware middleware [17] called Maestro, designed for data orchestration. In particular, we put the emphasis on the data abstraction layer and the API we have devised. We evaluate our approach with benchmarks and we discuss how Maestro can be of benefit for real-life use-cases such as numerical weather forecasting.

8.2   Cloud and Edge processing

8.2.1   Exploring the Computing Continuum through repeatable, replicable and reproducible Edge-to-Cloud experiments

Participants   Alexandru Costan, Gabriel Antoniu.

Collaboration.   *This work has been carried out in close co-operation with Pedro De Souza Bento Da Silva, formerly a post-doc student in the team, and now at Hasso Plattner Institute, Berlin, Germany.*

Distributed digital infrastructures for computation and analytics are now evolving towards an interconnected ecosystem allowing complex applications to be executed from IoT Edge devices to the HPC Cloud (aka the *Computing Continuum*, the *Digital Continuum*, or the *Transcontinuum*). Understanding end-to-end performance in such a complex continuum is challenging. This breaks down to reconciling many, typically contradicting application requirements and constraints with low-level infrastructure design choices. One important challenge is to accurately reproduce relevant behaviors of a given application workflow and representative settings of the physical infrastructure underlying this complex continuum.

In this work we introduce a rigorous methodology for such a process and validate it through E2Clab [16]. It is the first platform to support the complete analysis cycle of an application on the Computing Continuum: *(i)* the configuration of the experimental environment, libraries and frameworks; *(ii)* the mapping between the application parts and machines on the Edge, Fog and Cloud; *(iii)* the deployment of the application on the infrastructure; *(iv)* the automated execution; and *(v)* the gathering of experiment metrics. We illustrate its usage with a real-life application deployed on the Grid'5000 testbed, showing that our framework allows one to understand and improve performance, by correlating it to the parameter settings, the resource usage and the specifics of the underlying infrastructure.

8.2.2   Analytical models for performance evaluation of stream processing

Participants   Alexandru Costan, Gabriel Antoniu.

Collaboration.   *This work has been carried out in close co-operation with José Aguilar Canepa, formerly a PhD intern in the team, and now at Instituto Politécnico Nacionál, Mexico.*

A major challenge in deploying stream processing on Cloud and Edge infrastructure is finding the appropriate mapping between the application graph and the network graph. Operators need to be placed on machines in a way to enhance performance, i.e., reduce the total execution time. We are studying this challenge in the context of the SmartFastData associate team.

We have used a dynamic programming (DP) approach to address this mapping algorithm. The critical aspect of DP is to divide the problem into smaller sub-problems as independent as possible, that can be solved more easily than the original one, so that the sub-solutions can be combined to solve the complete problem. With this principle in mind, we designed an algorithm to solve the

Graph Coloring Problem, a classical combinatorial optimization problem. Given a graph, the goal is to assign each of its vertices a label (color) so that no adjacent vertices share the same label. In our case, colors correspond to different machines (associated with one operator).

Our proposal uses random, efficient algorithms to partition the graph into smaller sub-graphs, color them using a classical genetic algorithm, and glue together to solve the complete graph. Experimental tests on benchmark graphs show promising results (a paper is currently underway to showcase these initial results).

We are now interested in testing its performance on large graphs derived from real-life scenarios. A case study is a network of thousands of servers used to distribute contents on the Internet. This scenario is a typical scheduling application of a graph coloring problem.

8.2.3   Modeling smart cities applications

Participants   Alexandru Costan, Gabriel Antoniu.

Collaboration. *This work has been carried out in close co-operation with Edgar Romo Montiel, formerly a PhD intern in the team, and now at Instituto Politécnico Nacionál, Mexico.*

Some important components of smart cities are the connected vehicles. We have focused on modeling data flow from Vehicular Networks through phase-type distributions and Artificial Neural Networks to obtain a robust system to reserve and allocate resources needed to process the data collected from those vehicles. The ultimate goal is to distribute the data flow's prediction in multiple antennas to estimate load work in a specific point based on the demand for resources in the neighborhood.

To this end, we have leveraged the SUMO simulator that generates vehicles' traffic flow to feed the Neural Network and the Phase-type model. The first approach considered a single antenna, where the mathematical model was computing an upper-bound of the real traffic and the Neural Network was learning the actual traffic history. This single antenna model only considers its own data. Multiple antennas' model considers the data from neighbor antennas to feed the Neural Network to predict the vehicles' flow.

The precision of Neural Networks is beneficial for the prediction. However, the complexity of such a model increases so that updates or computing may require large Cloud capabilities. We leveraged Grid'5000 to compare performance between Cloud capabilities and Edge capabilities and estimate multiple antennas' data flow.

The resulting flow prediction models are currently in submission to a journal in the field of Vehicular/Edge Networks. This collaboration of Inria and Instituto Politécnico Nacional, Mexico (IPN) is the starting point of a future papers' collection for Edge/Cloud computing architecture in the area of Vehicular Networks.

8.2.4   Efficient InfiniBand access for stream ingestion systems

Participants   Thomas Bouvier, Alexandru Costan, Gabriel Antoniu.

Large-scale applications like Big Data analytics need efficient communication for their processing tasks. Fast networks like InfiniBand offer bandwidths of up to 200 Gb/s with latencies of less than two microseconds. While they are mainly used in high performance computing, some applications in the field of Big Data analytics are now starting to use them. In addition, some cloud

providers are offering instances equipped with InfiniBand hardware. Yet, the support for such fast networks is still limited in current stream processing engines.

To this end we propose to enhance KerA, our flagship stream ingestion system, with support for InfiniBand access. The idea is to allow the distribution and the access to the stream partitions and their corresponding metadata through InfiniBand. We leveraged the Neutrino network library developed by our partners from the University of Dusseldorf in the context of the PHC PROCOPE FlexStream project.

This library provides comfortable and efficient access to InfiniBand hardware in Java as well as a poll-based multithreaded connection management. Neutrino supports InfiniBand message passing as well as remote direct memory access, is implemented using the Java Native Interface, and can be used with any Java Virtual Machine. It also provides access to native C structures via a specific proxy system, which in turn enables the developer to leverage the full functionality of InfiniBand hardware.

Our initial experiments showed that efficient access to InfiniBand hardware is possible while fully utilizing the available bandwidth and thus speeding-up the stream ingestion.

8.3 Supporting AI across the digital continuum

Participants   Alexandru Costan, Gabriel Antoniu.

Collaboration. *This work has been carried out in close co-operation with Pedro De Souza Bento Da Silva, formerly a post-doc student in the team, and now at Hasso Plattner Institute, Berlin, Germany.*

The growth of the Internet of Things led to an explosion of data volumes at the Edge of the Internet. To reduce costs induced by data movement and centralized cloud-based processing, it becomes more and more relevant to process and analyze such data closer to the data sources. In particular, more and more applications are relying on decentralized Machine Learning (ML) applied to stream data.

However, exploiting Edge computing capabilities for stream-based processing is challenging. It requires to cope with complex characteristics and constraints imposed by all resources along the data path, and with a large set of heterogeneous data processing and management frameworks. The community needs tools to enable the modeling of that complexity and to integrate the various components involved. Furthermore, it would benefit from tools facilitating experimentation and understanding the performance of an application with respect to a series of implementation and deployment options.

We designed IntelliEdge, a hierarchical, layer-based approach for modeling distributed, stream-based applications on Edge-to-Cloud continuum infrastructures. The objective of IntelliEdge is to support integration of complex stream-based analytics applications atop the Edge-to-Cloud Continuum.

We demonstrated how IntelliEdge can be applied to a concrete real-life ML-based application — early earthquake warning — to help answer questions like: When is it worth decentralizing the classification load from the Cloud to the Edge and how?

9 Partnerships and cooperations

9.1 International initiatives

9.1.1 Inria International Labs

UNIFY

Title: *Intelligent Unified Data Services for Hybrid Workflows Combining Compute-Intensive Simulations and Data-Intensive Analytics at Extreme Scales*
Duration: 2019 - 2021
Coordinator: Gabriel Antoniu
Partners:

- Dpt of Mathematics - symbolic computation group, Argonne National Laboratory (United States)

Inria contact: Gabriel Antoniu
Summary: The landscape of scientific computing is being radically reshaped by the explosive growth in the number and power of digital data generators, ranging from major scientific instruments to the Internet of Things (IoT) and the unprecedented volume and diversity of the data they generate. This requires a rich, extended ecosystem including simulation, data analytics, and learning applications, each with distinct data management and analysis needs. Science activities are beginning to combine these techniques in new, large-scale workflows, in which scientific data is produced, consumed, and analyzed across multiple distinct steps that span computing resources, software frameworks, and time. This paradigm introduces new data-related challenges at several levels.

The UNIFY Associate Team aims to address three such challenges. First, to allow scientists to obtain fast, real-time insights from complex workflows combining extreme-scale computations with data analytics, we will explore how recently emerged Big Data processing techniques (e.g., based on stream processing) can be leveraged with modern in situ/in transit processing approaches used in HPC environments. Second, we will investigate how to use transient storage systems to enable efficient, dynamic data management for hybrid workflows combining simulations and analytics. Finally, the explosion of learning and AI provides new tools that can enable much more adaptable resource management and data services than available today, which can further optimize such data processing workflows.

As a result of collaborative work initiated during the previous years, two new joint papers were published in 2020 in the area of elastic distributed storage systems [14, 12].

At the 11th JLESC virtual workshop held in September 2020, Daniel Rosendo and François Tessier made presentations on subjects that are now in discussion for future collaborations inside the UNIFY Team. As a result, Daniel started to explore the usage of DeepHyper developed at ANL in conjunction with the E2Clab deployment framework for reproducible experimentation across the digital continuum developed by KerData, Inria. Goal: enable ML-based experimentation across the digital continuum.

In November 2020, Hugo Chaugier, Alexandu Costan, Gabriel Antoniu (Inria) and Bogdan Nicolae (ANL) started weekly discussions for a collaboration on the topic of parallel and distributed Deep Learning. Hugo Chaugier will start a MS internship on this topic in 2021.

9.1.2 Inria associate team not involved in an IIL

SmartFastData

Title: Efficient Data Management in Support of Hybrid Edge/Cloud Analytics for Smart Cities
Duration: 2019 – 2022

Coordinator: Alexandru Costan
Partners:

- Centro de Investigación en Computación, Instituto Politécnico Nacional (Mexico)

Inria contact: Alexandru Costan
Website: https://team.inria.fr/smartfastdata/
Summary: The proliferation of small sensors and devices that are capable of generating valuable information in the context of the Internet of Things (IoT) has exacerbated the amount of data flowing from all connected objects to private and public cloud infrastructures.
 In particular, this is true for Smart City applications, which cover a large spectrum of needs in public safety, water and energy management. Unfortunately, the lack of a scalable data management subsystem is becoming an important bottleneck for such applications, as it increases the gap between their I/O requirements and the storage performance.
 The vision underlying the SmartFastData associated team is that, by smartly and efficiently combining the data-driven analytics at the edge and in the cloud, it becomes possible to make a substantial step beyond state-of-the-art prescriptive analytics through a new, high-potential, faster approach to react to the sensed data.
 The goal is to build a data management platform that will enable comprehensive joint analytics of past (historical) and present (real-time) data, in the cloud and at the edge, respectively, allowing to quickly detect and react to special conditions and to predict how the targeted system would behave in critical situations.

In 2020, we leveraged the analytical models studied by the Instituto Politécnico Nacional (IPN) for hybrid deployments in order to propose a rigorous methodology for designing experiments with real-world workloads on the Computing Continuum spanning from the Edge through the Fog to the Cloud.

We have fine-tuned and optimized the Edge/Cloud mapping algorithms proposed earlier with the goal to incorporate Dynamic Programming's main principle into the internal engine of Genetic Algorithms, a popular technique used for numerical optimization, but seen as a poor choice when it comes to combinatorial optimization.

We modeled the data flow from Vehicular Networks through phase-type distributions and Artificial Neural Networks to obtain a robust system to reserve and allocate their needed computing resources.

We have further distributed the data flow's prediction from a single antenna to multiple antennas, which allows to estimate load work in a specific point based on the demand for resources in the neighborhood.

9.1.3 Participation in other international programs

FlexStream: Automatic elasticity for stream-based applications

Program: PHC PROCOPE 2020
Project acronym: FlexStream
Project title: Automatic Elasticity for Stream-based Applications
Duration: January 2020–December 2021
Coordinator: Alexandru Costan
Other partners: University of Dusseldorf (UDUS)
Summary: This project aims at developing concepts providing automatic scaling for stream processing applications. In particular, FlexStream aims at developing and evaluating a prototype which will integrate a stream ingestion-system from IRISA and an in-memory storage from UDUS.
 For this approach a tight cooperation is mandatory in order to be successful which in turn requires visits on both sides and longer exchanges, especially for the involved PhD students,

in order to allow an efficient integrated software design, development as well as joint experiments on large platforms and preparing joint publications.

In 2020, we focused on the aspects of efficient communication through InfiniBand (leveraging UDUS expertise in this domain) for distributed storage and ingestion frameworks like KerA, developed by KerData. During the visit of Alexandru Costan at UDUS the roadmap towards this goal was laid out as well as a plan for future mobilities between our two teams.

9.2 International research visitors

9.2.1 Visits to international teams

Research stays abroad

Alexandru Costan visited the team of Michael Schottner at University of Dusseldorf from March 3 to March 7, 2020, in the context of the PHC PROCOPE FlexStream project. Working closely with the UDUS team, he defined the work program for the upcoming year with respect to the project's objectives.

9.3 European initiatives

9.3.1 FP7 & H2020 Projects

Two new H2020 EuroHPC project proposals defined in 2020

> **Participants** Gabriel Antoniu, Alexandru Costan, Joshua Bowden, François Tessier.

In the area of HPC-Big Data-AI convergence, a major achievement in 2020 is the acceptance of *a new European EuroHPC project* called ACROSS (2021–2023). Gabriel Antoniu is coordinating Inria's contribution to this project, which includes a transfer action for the Damaris technology. In ACROSS, Damaris will be used by an application in the area of oil and gas, as part of a representative use case combining HPC, Big Data and AI.

Damaris was also selected as one of the two Inria software contributions to the software stack proposed for the future European Exascale machine (EUPEX EuroHPC project proposal, submitted in September 2020, planned to start in 2021). The EUPEX consortium aims to design, build, and validate the first EU platform for HPC, covering end-to-end the spectrum of required technologies with European assets: from the architecture, processor, system software, development tools to the applications.

PRACE 6th Implementation Phase Project (PRACE6-IP)

> **Participants** Joshua Bowden, Gabriel Antoniu.

Program: H2020 Research and Innovation Action (RIA), call H2020-INFRAEDI-2018-1
Project acronym: PRACE-6IP
Project title: PRACE 6th Implementation Phase Project
Duration: May 2019–Dec 2021
Coordinator: FZJ

Other partners: HLRS, LRZ, GENCI, CEA, CINES, CNRS, IDRIS, Inria, EPCC, BSC, CESGA, CSC, ETH-CSCS, SURFsara, KTH-SNIC, CINECA, PSNC, CYFRONET, WCNS, UiOsingma2, GRNET, UC-LCA, Univ MINHO, ICHEC, UHEM, CASTORCm NCSA, IT4I-VSB, KIFU, UL, CCSAS, CENAERO, Univ Lux, GEANT

Web site: https://cordis.europa.eu/project/id/823767

PRACE, the Partnership for Advanced Computing is the permanent pan-European High Performance Computing service providing world-class systems for world-class science. Systems at the highest performance level (Tier-0) are deployed by Germany, France, Italy, Spain and Switzerland, providing researchers with more than 17 billion core hours of compute time. HPC experts from 25 member states enable users from academia and industry to ascertain leadership and remain competitive in the Global Race. Currently PRACE is finalizing the transition to PRACE 2, the successor of the initial five-year period.

The objectives of PRACE-6IP are to build on and seamlessly continue the success of PRACE and start new innovative and collaborative activities proposed by the consortium. These include: assisting the development of PRACE 2; strengthening the internationally recognized PRACE brand; continuing and extend advanced training which so far provided more than 36,400 person-training days; preparing strategies and best practices towards Exascale computing, work on forward-looking SW solutions; coordinating and enhancing the operation of the multi-tier HPC systems and services; and supporting users to exploit massively parallel systems and novel architectures. A high-level Service Catalog is provided.

The proven project structure will be used to achieve each of the objectives in 7 dedicated work packages. The activities are designed to increase Europe's research and innovation potential especially through: seamless and efficient Tier-0 services and a pan-European HPC ecosystem including national capabilities; promoting take-up by industry and new communities and special offers to SMEs; assistance to PRACE 2 development; proposing strategies for deployment of leadership systems; collaborating with the ETP4HPC, CoEs and other European and international organizations on future architectures, training, application support and policies. This will be monitored through a set of KPIs.

In PRACE-6IP, the Damaris framework developed by the KerData team is being experimented as a technology to provide a service for in situ visualization and processing to PRACE users. In this context, a demonstrator is currently being built using Damaris for Code_Saturne, a CFD application.

9.3.2   Collaborations in European programs, except FP7 and H2020

The ENGAGE Inria-DFKI project proposal

> **Participants**   Gabriel Antoniu, Alexandru Costan, Thomas Bouvier, Daniel Rosendo.

In the area of HPC-AI convergence, Gabriel Antoniu coordinates the ENGAGE Inria-DFKI project proposal (submitted, under evaluation). In addition to the KerData team, it involves two other Inria teams: HiePACS (Bordeaux) and DATAMOVE (Grenoble). It aims to create foundations for a new generation of high-performance computing (HPC) environments for Artificial Intelligence (AI) workloads.

The basic premise for these workloads is that in the future, training data for Deep Neural Networks (DNN) will no longer only be stored and processed in epochs, but rather be generated on-the-fly using parametric models and simulations. This is particularly useful in situations where obtaining data by other means is expensive or difficult or where a phenomenon has been predicted in theory, but not yet observed. One key application of this approach is to validate and certify AI systems through targeted testing with synthetically generated data from simulations.

The project proposes contributions on three levels: On the application level, it will address the question how the adaptive sampling of parameter spaces will allow for better choices on what data to generate. On the middleware level, it will address the question how virtualization and scheduling need to be adapted to facilitate and optimize the execution of resulting mixed workloads consisting of training and simulation tasks, running on potentially hybrid (HPC/cloud/edge) infrastructures. On the resource management level, it will contribute to novel strategies to optimize memory management and dynamic selection of parallel resources to run the training and inference phases.

In summary, the project will create a blueprint for a new generation of AI compute infrastructures that goes beyond the concept of epoch-based data management and considers model-based online-training of Neural Networks as the new paradigm for DNN applications.

9.3.3   Collaborations with major European organizations

Participants   Gabriel Antoniu, Alexandru Costan.

Appointments by Inria in relation to European bodies

Big Data Value Association (BDVA) and ETP4HPC: Since 2018, Gabriel Antoniu and Alexandru Costan have been serving as Inria representatives in the working groups dedicated to *HPC-Big Data* convergence.

Community service at European level in response to external invitations

ETP4HPC: Since 2019, Gabriel Antoniu has served as a co-leader of the working group on Programming Environments and co-leader of two research clusters, contributing to the Strategic Research Agenda of ETP4HPC, published in March 2020. Alexandru Costan served as a member of these working groups. Activity is now continuing through regular meetings to refine the ETP4HPC technical focus areas to prepare the future edition of the SRA that should be published in 2022.

Transcontinuum Initiative (TCI): In 2020, as a follow-up action to the publication of its Strategic Research Agenda, ETP4HPC initiated a collaborative initiative called TCI (Transcontinuum Initiative). It gathers major European associations in the areas of HPC, Big Data, AI, 5G, Cybersecurity, including ETP4HPC, BDVA, CLAIRE, HIPEAC, 5G IA, ECSO). It aims to strengthen research and industry in Europe to support the Digital Continuum - infrastructure (including HPC systems, clouds, edge infrastructures) by helping to define a set of research focus areas/topics requiring interdisciplinary action.
The expected outcome of this effort is the co-editing of multidisciplinary calls for projects to be funded by the European Commission. Gabriel Antoniu is in charge of ensuring the BDVA-ETP4HPC coordination and of co-animating the working group dedicated to the definition of representative application use cases.

Big Data Value Association: In 2020, Gabriel Antoniu was asked by BDVA to start coordinating BDVA's contribution to the TCI initiative recently started (see above). He also participated to the organization of a joint BDVA-ETP4HPC seminar on HPC, Big Data, IoT and AI future industry-driven collaborative strategic topics.

ZettaFlow: Unified fast-data storage and analytics platform for IoT

Participants   Gabriel Antoniu, Alexandru Costan.

Program: EIT Digital Innovation Factory
Project acronym: ZettaFlow
Project title: ZettaFlow: Unified Fast Data Storage and Analytics Platform for IoT
Duration: October 2019–December 2020
Technical Coordinator: Ovidiu Marcu
Other partners: Technische Universität Berlin and System@tic
Web site: https://zettaflow.io/

The objective of this project was to create a startup in order to commercialize the ZettaFlow platform: a dynamic, unified and auto-balanced real-time storage and analytics industrial IoT platform. ZettaFlow was based on KerA, a streaming storage system prototype developed within the KerData team.

The project was stopped by the end of 2020 due to the departure of Ovidiu Marcu, the expected CTO of the targeted startup, and to difficulties in hiring in 2020, at the beginning of the Covid pandemic.

9.4 National initiatives

9.4.1 ANR
OverFlow (2015–2021)

Participants Alexandru Costan, Daniel Rosendo, Gabriel Antoniu.

Project Acronym: OverFlow
Project Title: Workflow Data Management as a Service for Multisite Applications
Coordinator: Alexandru Costan
Duration: October 2015–March 2021
Other Partners: None (Young Researcher Project, JCJC)
External collaborators: Kate Keahey (University of Chicago and Argonne National Laboratory), Bogdan Nicolae (Argonne National Lab)
Web site: https://sites.google.com/view/anroverflow

This project investigates approaches to data management enabling an efficient execution of geographically distributed workflows running on multi-site clouds.

In 2020, we focused on the reproducibility aspects of the workflows deployed on hybrid edge-cloud infrastructures. This involves the non trivial task of reconciling many, typically contradicting application requirements and constraints with low-level infrastructure design choices.

This year we introduced a rigorous methodology for such a process. It allows one to understand and improve performance, by correlating it to the parameter settings, the resource usage and the specifics of the underlying infrastructure. Eventually, this methodology allows other researchers to leverage the experimental results and advance knowledge in different domains, by enabling three R's of research quality: Repeatability, Replicability, and Reproducibility.

9.4.2 Other National Projects
HPC-Big Data Inria Inria Challenge (ex-IPL)

Participants Daniel Rosendo, Gabriel Antoniu, Alexandru Costan.

Collaboration. *This work has been carried out in close co-operation with Pedro De Souza Bento Da Silva, formerly a post-doc student in the team, and now at Hasso Plattner Institute, Berlin, Germany.*

Project Acronym: HPC-BigData
Project Title: The HPC-BigData INRIA Challenge
Coordinator: Bruno Raffin
Duration: 2018–2022
Web site: https://project.inria.fr/hpcbigdata/

The goal of this HPC-BigData IPL is to gather teams from the HPC, Big Data and Machine Learning (ML) areas to work at the intersection between these domains. Research is organized along three main axes: high performance analytics for scientific computing applications, high performance analytics for big data applications, infrastructure and resource management. Gabriel Antoniu is a member of the Advisory Board and leader of the Frameworks work package.

In 2020 we contributed with the proposal of the E2Clab methodology and supporting framework for Edge-to-Cloud Experimentation [16] (see details in Section 8.2).

ADT Damaris 2

Participants	Joshua Charles Bowden, Gabriel Antoniu.

Project Acronym: ADT Damaris 2
Project Title: Technology development action for the Damaris environment
Coordinator: Gabriel Antoniu
Duration: 2019–2022
Web site: https://project.inria.fr/damaris/

This action aims to support the development of the Damaris software. Inria's *Technological Development Office* (D2T, *Direction du Développement Technologique*) provided 3 years of funding support for a senior engineer.

In April 2020, Joshua Bowden was hired on this position. He introduced a support for unstructured mesh model types to Damaris. This capability opens up the use of Damaris for a large number of simulation types that depend on this data structure, in the areas of Computational Fluid dynamics, which has applications in energy production and combustion modeling, electric modeling and atmospheric and flow.

The capability has been developed and tested using Code_Saturne, a finite volume computational fluid dynamics (CFD) simulation environment. Code_Saturne is an open source y CFD modeling environment which supports both single phase and multi-phase flow and includes modules for atmospheric flow, combustion modeling, electric modeling and particle tracking. This work is being validated on PRACE Tier-0 computing infrastructure in the framework of the PRACE-6IP project.

Grid'5000   We are members of Grid'5000 community and run experiments on the Grid'5000 platform on a daily basis.

10   Dissemination

10.1   Promoting scientific activities

10.1.1   Scientific events: organization

General chair, scientific chair

Luc Bougé: Steering Committee Chair of the Euro-Par International Conference on Parallel and
Distributed Computing. Euro-Par celebrated its 25th anniversary in Göttingen, Germany, in
2019.

François Tessier: Co-Chair of SuperCompCloud, the 3rd Workshop on Interoperability of Super-
computing and Cloud Technologies held in conjunction with Supercomputing 20.

Member of the organizing committees

François Tessier: Member of the organizing committee of SuperCompCloud, the 3rd Workshop
on Interoperability of Supercomputing and Cloud Technologies held in conjunction with
Supercomputing 20.

10.1.2   Scientific events: selection

Chair of conference program committees

Gabriel Antoniu: Program Chair of HPS'20 — 1st Workshop on High-Performance Storage, held
in conjunction with the IEEE IPDPS 2020 conference.

Member of the conference program committees

Alexandru Costan: IEEE/ACM SC'20 (Posters and ACM Student Research Competition), IEEE
Cluster 2020, IEEE/ACM UCC 2020, IEEE Big Data 2020, IEEE CloudCom 2020.
Gabriel Antoniu: IEEE IPDPS 2020, Euro-Par 2020, SuperCompCloud 2020.
François Tessier: IEEE/ACM CCGrid 2020.

Reviewer

Alexandru Costan: ACM HPDC 2020, IEEE IPDPS 2020, IEEE/ACM CCGrid 2020.

10.1.3   Journal

Member of the editorial boards

Gabriel Antoniu: Associate Editor of JPDC - the Elsevier Journal of Parallel and Distributed
Computing.

Reviewer - reviewing activities

Alexandru Costan: IEEE Transactions on Parallel and Distributed Systems, Future Generation
Computer Systems, Concurrency and Computation Practice and Experience, IEEE Transac-
tions on Cloud Computing, Journal of Parallel and Distributed Computing.
Gabriel Antoniu: IEEE Transactions on Parallel and Distributed Systems, SoftwareX.

10.1.4 Invited talks

Gabriel Antoniu:

- LIRIMA seminar, in May 2020. *The Edge, the Cloud and the Supercomputer: Welcome to the Age of the Digital Continuum!*. URL: https://videos-rennes.inria.fr/video/rJgGW8ATU.
- BDVA Activity Group, November 2020. *TCI: The Transcontinuum Initiative*. URL: https://www.bdva.eu/node/1680.

10.1.5 Leadership within the scientific community

Gabriel Antoniu:

TCI: Since 2020, co-leader of the Use-Case Analysis Working Group. TCI (The Transcontinuum Initiative) emerged as a collaborative initiative of ETP4HPC, BDVA, CLAIRE and other peer organizations, aiming to identify joint research challenges for leveraging the HPC-Cloud-Edge computing continuum and make recommendations to the European Commission about topics to be funded in upcoming calls for projects.

ETP4HPC: Since 2019, co-leader of the working group on Programming Environments and co-lead of two research clusters, contributing to the next Strategic Research Agenda of ETP4HPC (published in March 2020).

International lab management: *Vice Executive Director of JLESC* for Inria. JLESC is the Joint Inria-Illinois-ANL-BSC-JSC-RIKEN/AICS Laboratory for Extreme-Scale Computing. Within JLESC, he also serves as a *Topic Leader* for Data storage, I/O and in situ processing for Inria.

Team management: *Head of the KerData Project-Team* (INRIA-ENS Rennes-INSA Rennes).

International Associate Team management: Leader of the UNIFY Associate Team with Argonne National Lab (2019–2021).

Technology development project management: Coordinator of the Damaris 2 ADT project (2019–2022).

Luc Bougé:

SIF: Co-Vice-President of the *French Society for Informatics* (*Société informatique de France*, SIF), in charge of the Teaching Department.

CoSSAF: Member of the Foundation Committee for the *College of the French Scientific Societies* (CoSSAF, *Collège des sociétés savantes académiques de France*) gathering more than 40 French societies from all domains.

Alexandru Costan:

International Associate Team management: Leader of the SmartFastData Associate Team with Instituto Politécnico Nacional, Mexico City (2019–2021).

François Tessier:

JLESC: Member of the organizing committee of the bi-annual workshops of the Joint Laboratory for Extreme Scale Computing.

10.1.6 Teaching

Alexandru Costan

- Bachelor: Software Engineering and Java Programming, 28 hours (lab sessions), L3, INSA Rennes.
- Bachelor: Databases, 68 hours (lectures and lab sessions), L2, INSA Rennes, France.
- Bachelor: Practical case studies, 24 hours (project), L3, INSA Rennes.
- Master: Big Data Storage and Processing, 28h hours (lectures, lab sessions), M1, INSA Rennes.

- Master: Algorithms for Big Data, 28 hours (lectures, lab sessions), M2, INSA Rennes.
- Master: Big Data Project, 28 hours (project), M2, INSA Rennes.

Gabriel Antoniu

- Master (Engineering Degree, 5th year): Big Data, 24 hours (lectures), M2 level, ENSAI (*École nationale supérieure de la statistique et de l'analyse de l'information*), Bruz, France.
- Master: Scalable Distributed Systems, 10 hours (lectures), M1 level, SDS Module, EIT ICT Labs Master School, France.
- Master: Infrastructures for Big Data, 10 hours (lectures), M2 level, IBD Module, SIF Master Program, University of Rennes, France.
- Master: Cloud Computing and Big Data, 14 hours (lectures), M2 level, Cloud Module, MIAGE Master Program, University of Rennes, France.

Daniel Rosendo

- Master: Miage BDDA, 24 hours (lab sessions), M2, ISTIC Rennes.
- Bachelor: Algorithms for Big Data, 7.5 hours (lectures, lab sessions), INSA Rennes.

10.1.7 Supervision

- PhD in progress : Daniel Rosendo *Enabling HPC-Big Data Convergence for Intelligent Extreme-Scale Analytics*, thesis started in October 2019, co-advised by Gabriel Antoniu, Alexandru Costan and Patrick Valduriez (Inria).

10.1.8 Juries

Luc Bougé: Member of the jury the *CAPES of mathématiques*, *Informatics track*. This national committee selects more than 1000 mathematics teachers per year for French secondary schools and high-schools.

Gabriel Antoniu: Referee for 3 PhD juries for PhD defenses at Barcelona Supercomputing Center, École Normale Supérieure de Lyon, University of Montpellier. Member of a PhD jury at the University of Bordeaux.

10.2 Popularization

10.2.1 Internal or external Inria responsibilities

Gabriel Antoniu:

ETP4HPC and BDVA: Inria representative in the working groups of BDVA and ETP4HPC dedicated to HPC-Big Data convergence.

Luc Bougé:

ANR: Scientific officer at the *National Research Agency* (ANR, *Agence nationale de la recherche*) in charge of data management.

PNRIA: Porgramme National de Recherche en Intelligence Artificielle (*National Research Program for Artificial Intelligence*). Member of the management team for the IA for Humanity governmental program launched in March 2018. Bertrand Braunschweig, Inria, and now Isabelle Herlin, are the scientific directors of the program.

Alexandru Costan:

In charge of internships at the Computer Science Department of INSA Rennes.

In charge of the organization of the IRISA D1 Department Seminars.

In charge of the management of the KerData team access to Grid'5000.

10.2.2 Interventions

Luc Bougé:

IH2EF: Participation to a series of trainings at the Institut des hautes études de l'éducation et de la formation (IH2EF, *Institute for Higher Studies in Education and Training*) about Artificial Intelligence. It targeted the persons in charge of pedagogy high-schools all over France.

ANR: Organization of various seminars and training about Large-Scale Data Management at *National Research Agency* (ANR, *Agence nationale de la recherche*). It focused on using Web API to access databases, for instance HAL and scanR, the search engine for French research and innovation.

11 Scientific production

11.1 Major publications

[1] N. Cheriere, M. Dorier and G. Antoniu. 'How fast can one resize a distributed file system?' In: *Journal of Parallel and Distributed Computing* 140 (June 2020), pp. 80–98. DOI: 10.1016/j.jpdc.2020.02.001. URL: https://hal.archives-ouvertes.fr/hal-02961875.

[2] M. Dorier, G. Antoniu, F. Cappello, M. Snir, R. Sisneros, O. Yildiz, S. Ibrahim, T. Peterka and L. Orf. 'Damaris: Addressing Performance Variability in Data Management for Post-Petascale Simulations'. In: *ACM Transactions on Parallel Computing* (2016). URL: https://hal.inria.fr/hal-01353890.

[3] M. Dorier, S. Ibrahim, G. Antoniu and R. Ross. 'Using Formal Grammars to Predict I/O Behaviors in HPC: the Omnisc'IO Approach'. In: *TPDS - IEEE Transactions on Parallel and Distributed Systems* (Oct. 2015). DOI: 10.1109/TPDS.2015.2485980. URL: https://hal.inria.fr/hal-01238103.

[4] K. Fauvel, D. Balouek-Thomert, D. Melgar, P. Silva, A. Simonet, G. Antoniu, A. Costan, V. Masson, M. Parashar, I. Rodero and A. Termier. 'A Distributed Multi-Sensor Machine Learning Approach to Earthquake Early Warning'. In: *In Proceedings of the 34th AAAI Conference on Artificial Intelligence*. New York, United States, Feb. 2020, pp. 403–411. DOI: 10.1609/aaai.v34i01.5376. URL: https://hal.archives-ouvertes.fr/hal-02373429.

[5] J. Liu, L. Pineda, E. Pacitti, A. Costan, P. Valduriez, G. Antoniu and M. Mattoso. 'Efficient Scheduling of Scientific Workflows using Hot Metadata in a Multisite Cloud'. In: *IEEE Transactions on Knowledge and Data Engineering* (2018). DOI: 10.1109/TKDE.2018.2867857. URL: https://hal-lirmm.ccsd.cnrs.fr/lirmm-01867717.

[6] O.-C. Marcu, A. Costan, G. Antoniu, M. S. Pérez-Hernández, B. Nicolae, R. Tudoran and S. Bortoli. 'KerA: Scalable Data Ingestion for Stream Processing'. In: *ICDCS 2018 - 38th IEEE International Conference on Distributed Computing Systems*. Vienna, Austria: IEEE, July 2018, pp. 1480–1485. DOI: 10.1109/ICDCS.2018.00152. URL: https://hal.inria.fr/hal-01773799.

[7] O.-C. Marcu, A. Costan, G. Antoniu and M. S. Pérez-Hernández. 'Spark versus Flink: Understanding Performance in Big Data Analytics Frameworks'. In: *Cluster 2016 - The IEEE 2016 International Conference on Cluster Computing*. Taipei, Taiwan, Sept. 2016. URL: https://hal.inria.fr/hal-01347638.

[8] P. Matri, Y. Alforov, A. Brandon, M. Pérez, A. Costan, G. Antoniu, M. Kuhn, P. Carns and T. Ludwig. 'Mission Possible: Unify HPC and Big Data Stacks Towards Application-Defined Blobs at the Storage Layer'. In: *Future Generation Computer Systems* 109 (Aug. 2020), pp. 668–677. DOI: 10.1016/j.future.2018.07.035. URL: https://hal.archives-ouvertes.fr/hal-01892682.

[9] P. Matri, A. Costan, G. Antoniu, J. Montes and M. S. Pérez. 'Týr: Blob Storage Meets Built-In Transactions'. In: *IEEE ACM SC16 - The International Conference for High Performance Computing, Networking, Storage and Analysis 2016*. Salt Lake City, United States, Nov. 2016. URL: https://hal.inria.fr/hal-01347652.

[10] D. Rosendo, P. Silva, M. Simonin, A. Costan and G. Antoniu. 'E2Clab: Exploring the Computing Continuum through Repeatable, Replicable and Reproducible Edge-to-Cloud Experiments'. In: *Cluster 2020 - IEEE International Conference on Cluster Computing*. Kobe, Japan, Sept. 2020, pp. 1–11. DOI: 10.1109/CLUSTER49012.2020.00028. URL: https://hal.archives-ouvertes.fr/hal-02916032.

[11] Y. Taleb, R. Stutsman, G. Antoniu and T. Cortes. 'Tailwind: Fast and Atomic RDMA-based Replication'. In: *ATC '18 - USENIX Annual Technical Conference*. Boston, United States, July 2018, pp. 850–863. URL: https://hal.inria.fr/hal-01676502.

11.2   Publications of the year

International journals

[12] N. Cheriere, M. Dorier and G. Antoniu. 'How fast can one resize a distributed file system?' In: *Journal of Parallel and Distributed Computing* 140 (June 2020), pp. 80–98. DOI: 10.1016/j.jpdc.2020.02.001. URL: https://hal.archives-ouvertes.fr/hal-02961875.

[13] P. Matri, Y. Alforov, A. Brandon, M. Pérez, A. Costan, G. Antoniu, M. Kuhn, P. Carns and T. Ludwig. 'Mission Possible: Unify HPC and Big Data Stacks Towards Application-Defined Blobs at the Storage Layer'. In: *Future Generation Computer Systems* 109 (Aug. 2020), pp. 668–677. DOI: 10.1016/j.future.2018.07.035. URL: https://hal.archives-ouvertes.fr/hal-01892682.

International peer-reviewed conferences

[14] N. Cheriere, M. Dorier, G. Antoniu, S. M. Wild, S. Leyffer and R. Ross. 'Pufferscale: Rescaling HPC Data Services for High Energy Physics Applications'. In: CCGRID -2020 - 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing. 2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID). Melbourne, Australia, 11th May 2020, pp. 182–191. DOI: 10.1109/CCGrid49817.2020.00-75. URL: https://hal.archives-ouvertes.fr/hal-02961907.

[15] K. Fauvel, D. Balouek-Thomert, D. Melgar, P. Silva, A. Simonet, G. Antoniu, A. Costan, V. Masson, M. Parashar, I. Rodero and A. Termier. 'A Distributed Multi-Sensor Machine Learning Approach to Earthquake Early Warning'. In: In Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York, United States, 7th Feb. 2020, pp. 403–411. DOI: 10.1609/aaai.v34i01.5376. URL: https://hal.archives-ouvertes.fr/hal-02373429.

[16] D. Rosendo, P. Silva, M. Simonin, A. Costan and G. Antoniu. 'E2Clab: Exploring the Computing Continuum through Repeatable, Replicable and Reproducible Edge-to-Cloud Experiments'. In: Cluster 2020 - IEEE International Conference on Cluster Computing. Kobe, Japan, 14th Sept. 2020, pp. 1–11. DOI: 10.1109/CLUSTER49012.2020.00028. URL: https://hal.archives-ouvertes.fr/hal-02916032.

11.3   Cited publications

[17] C. Haine, U.-U. Haus, F. Tessier, D. Sármány, S. Smart, T. Quintino and A. Tate. 'A Middleware API for Data Movement in Scientific Workflows'. In: International Supercomputing Conference (ISC 2021). Paper submitted. Under review. Frankfurt, Germany, 23rd June 2021.

[18]   T. Akidau, A. Balikov, K. Bekiroğlu, S. Chernyak, J. Haberman, R. Lax, S. McVeety, D. Mills, P. Nordstrom and S. Whittle. 'MillWheel: fault-tolerant stream processing at internet scale'. In: *Proceedings of the VLDB Endowment* 6.11 (2013), pp. 1033–1044.

[19]   R. Bolze, F. Cappello, E. Caron, M. Dayde, F. Desprez, E. Jeannot, Y. Jégou, S. Lanteri, J. Leduc, N. Melab, G. Mornet, R. Namyst, P. Primet, B. Quétier, O. Richard, E.-G. Talbi and I. Touche. 'Grid'5000: A Large Scale And Highly Reconfigurable Experimental Grid Testbed'. In: *International Journal of High Performance Computing Applications* 20.4 (2006), pp. 481–494. DOI: 10.1177/1094342006070078. URL: https://hal.inria.fr/hal-00684943.

[20]   J. Dean and S. Ghemawat. 'MapReduce: simplified data processing on large clusters'. In: *Communications of the ACM* 51.1 (2008), pp. 107–113.

[21]   A. Joly, P. Bonnet, H. Goëau, J. Barbe, S. Selmi, J. Champ, S. Dufour-Kowalski, A. Affouard, J. Carré, J.-F. o. Molino, N. Boujemaa and D. Barthélémy. 'A look inside the Pl@ntNet experience'. In: *Multimedia Systems* 22.6 (2016), pp. 751–766. DOI: 10.1007/s00530-015-046 2-9. URL: https://hal.inria.fr/hal-01182775.

[22]   *Amazon Elastic Map-Reduce (EMR)*. 2017. URL: https://aws.amazon.com/emr/.

[23]   *Chameleon Cloud*. 2021. URL: https://www.chameleoncloud.org/.

[24]   *Cybeletech - Digital technologies for the plant world*. 2021. URL: https://www.cybeletech.com/en/home/.

[25]   *Digital Single Market*. 2015. URL: https://ec.europa.eu/digital-single-market/en/digital-single-market/.

[26]   *ECMWF - European Centre for Medium-Range Weather Forecasts*. 2021. URL: https://www.ecmwf.int/.

[27]   *European Exascale Software Initiative*. 2013. URL: http://www.eesi-project.eu/.

[28]   *The European Technology Platform for High-Performance Computing*. 2012. URL: http://www.etp4hpc.eu/.

[29]   *European Cloud Strategy*. 2012. URL: https://ec.europa.eu/digital-single-market/en/european-cloud-computing-strategy/.

[30]   *Apache Flink*. 2016. URL: http://flink.apache.org/.

[31]   *International Exascale Software Program*. 2011. URL: http://www.exascale.org/iesp/.

[32]   *Inria's strategic plan "Towards Inria 2020"*. 2016. URL: https://www.inria.fr/recherche-innovation/.

[33]   *Apache Spark*. 2017. URL: http://spark.apache.org/.

[34]   *Storm*. 2014. URL: http://storm.apache.org/.